Subterranean Biology 56: 51–83 (2026)
doi: 10.3897/subtbiol.56.190752
https://subtbiol.pensoft.net

RESEARCH ARTICLE

A peer-reviewed open-access journal
Subterranean Biology
Published by The International Society for Subterranean Biology

# Nowhere to hide: Ensemble of Small Models exposes climate vulnerability and conservation gaps for an extremely rare troglobitic beetle genus

Zi-Wei Yin[1]

1 *Laboratory of Systematic Entomology, College of Life Sciences, Shanghai Normal University, Shanghai, China*

Corresponding author: Zi-Wei Yin (zwyin@shnu.edu.cn, pselaphinae@gmail.com)

Academic editor: Leonardo Latella | Received 6 March 2026 | Accepted 14 April 2026 | Published 23 April 2026

*https://zoobank.org/462C566A-AC1E-40F5-ADC0-78A570B6B606*

**Citation:** Yin Z-W (2026) Nowhere to hide: Ensemble of Small Models exposes climate vulnerability and conservation gaps for an extremely rare troglobitic beetle genus. Subterranean Biology 56: 51–83. https://doi.org/10.3897/subtbiol.56.190752

## Abstract

Subterranean ecosystems harbor a high diversity of narrow-range endemics, yet they remain largely invisible in conservation planning, representing a major bias that may trigger irreversible biodiversity loss under rapid climate change. This study applied an Ensemble of Small Models (ESMs) framework, coupled with a target-group background strategy, to assess the climate vulnerability and protection status of *Zopherobatrus* Yin & Li, 2015 (Coleoptera: Staphylinidae: Pselaphinae), an extremely rare troglobitic rove beetle genus endemic to the karst landscapes of southwest China. Habitat suitability was projected under current conditions and across the 2050s and 2070s (SSP1-2.6 and SSP5-8.5). Under current conditions, highly suitable habitat is strictly confined to specific karst terrains; future projections reveal that the retained habitat ratio declines to only 24.0–48.2%, while the conventional net area change metric (–36.4% to –63.1%) masks functional habitat loss because climatically "gained" areas remain ecologically inaccessible to this dispersal-limited group, and the projected range centroid shifts northwestward by up to 141 km. A systematic gap analysis, cross-referenced with national and provincial nature reserves, showed that only 2.6% of current suitable habitat is formally protected, all known localities fall entirely outside reserves, and 95.9% of identified climate refugia lack any legal protection. These findings quantified a "double jeopardy" for subterranean biodiversity: simultaneous habitat erosion driven by climate change combined with a near-total absence of formal protection. The establishment of targeted micro-reserves centered on climate refugia, the formal inclusion of *Zopherobatrus* in national protection lists, and the integration of subterranean fauna into regional conservation frameworks are recommended.

Copyright *Zi-Wei Yin*. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Abstract in Chinese**

地下生态系统孕育了丰富的狭域特有种，但在生物多样性保护规划中却长期处于被忽视的边缘地位；在气候快速变化背景下，这种认知偏差或将导致不可逆的生态后果。本研究采用小样本模型集成框架，结合目标类群背景策略，对特产于中国西南喀斯特地貌的罕见真洞穴蚁甲——幽蚁甲属*Zopherobatrus*的气候脆弱性及受保护现状进行了全面评估，并预测了其在当前及未来（2050年代和2070年代，基于SSP1-2.6 和SSP5-8.5 情景）的栖息地适宜性。结果表明，在当前气候条件下，该属的高度适宜栖息地严格局限于特定喀斯特生境。未来预测显示，其原有栖息地保留率将骤降至24.0%–48.2%。尽管模型预测出新的气候适宜区，但对于此类扩散能力受限的洞穴生物而言，这些“新增”区域在生态学上是不可达的，导致常规净面积变化指标（–36.4%至–63.1%）掩盖了其功能性栖息地丧失的严重程度；此外，预测显示其适生区质心将向西北方向发生最远达141公里的偏移。通过与国家级和省级自然保护区进行空间叠加的系统性保护空缺分析发现，当前仅有2.6%的适宜栖息地被纳入正式保护体系；所有已知分布点均完全游离于现有自然保护区之外；且高达95.9%的潜在气候避难所缺乏任何形式的法定保护。本研究定量揭示了地下生物多样性所面临的“双重危机”，即由气候变化驱动的栖息地持续退缩，以及几乎完全处于法定保护体系之外的严峻现状。鉴于此，本文呼吁建立以气候避难所为核心的针对性“微型保护区”，建议将幽蚁甲属正式列入国家重点保护野生动物名录，并亟需将地下动物群系全面纳入区域性生物多样性保护框架。

**Keywords**

Climate refugia, Ensemble of Small Models, gap analysis, karst, rove beetle, subterranean biodiversity conservation

## Introduction

Subterranean ecosystems are increasingly recognized as important reservoirs of biological diversity, harboring a disproportionate concentration of narrow-range endemic species whose evolutionary lineages have persisted in isolation for millions of years (Juberthie and Decu 1994; Culver and Pipan 2019; Mammola et al. 2019). Characterized by stringent environmental constraints, i.e., permanent darkness, high humidity, and well-buffered thermal regimes, these environments, particularly limestone caves within vast karst landscapes, have driven the evolution of highly specialized organisms exhibiting troglomorphic adaptations (e.g., depigmentation, anophthalmia, and elongated appendages). Despite their biological and eco-evolutionary significance, subterranean habitats are frequently relegated to the status of a “forgotten biome” in conservation biology. They are systematically overlooked in global and regional biodiversity assessments, largely due to logistical inaccessibility, taxonomic impediments, and a pervasive “out of sight, out of mind” bias (Cardoso et al. 2011; Ficetola et al. 2019; Mammola et al. 2019).

Although cave environments are traditionally perceived as highly buffered thermal refugia, this assumption is increasingly challenged by mounting evidence demonstrating that surface climate change readily permeates the subsurface through altered hydrological regimes, shifts in vegetation cover, and modified soil respiration (Mammola et al. 2018). However, the thermal stability of caves is relative, not absolute. While shallow subterranean zones remain tightly coupled to surface dynamics, deeper cave

environments exhibit significant thermal inertia; mean annual cave temperatures track surface trends, albeit with dampened amplitudes and temporal lags (Badino 2010). Nevertheless, at decadal scales, persistent surface warming is expected to equilibrate underground. This equilibration poses an existential threat to obligate cave-dwelling organisms, which, having evolved within narrow thermal envelopes, are expected to be disproportionately vulnerable to even modest warming (Mammola et al. 2018, 2019). Consequently, for these specialized taxa, the capacity for compensatory range shifts is further constrained by the discontinuous, island-like distribution of suitable subterranean habitats within highly fragmented karst landscapes (Christman et al. 2005).

The inadequacy of existing protected area (PA) networks in safeguarding subterranean biodiversity represents a parallel and compounding threat. Historically, the global PA estate has been designed with a primary focus on surface vegetation types, charismatic megafauna, and aboveground ecosystem services, with little consideration given to belowground habitats or the fauna they support (Mammola et al. 2019, 2020, 2024). This surface-centric approach implicitly assumes that the "umbrella effect" of terrestrial PAs extends downward into subsurface ecosystems, an assumption that remains largely untested and may be inadequate for taxa whose distributions are dictated by geological substrates and hydrogeological connectivity rather than surface land cover (Christman et al. 2005). In the diverse karst regions of southern China, which harbor one of the globe's richest assemblages of cave-adapted invertebrates (Deharveng and Bedos 2012; Tian et al. 2017, 2023), the extent to which existing nature reserves provide effective coverage for subterranean obligate species remains critical, yet largely unexamined.

The recently established genus *Zopherobatrus* Yin & Li, 2015 (Coleoptera: Staphylinidae) (Yin and Li 2015; Yin and He 2020; Yin 2020, 2025) serves as an ideal model system for investigating the intersection of climate vulnerability and conservation neglect in subterranean fauna. As an obligate cave-dwelling (troglobitic) lineage endemic to the karst terrains of southwest China, *Zopherobatrus* exhibits a full suite of troglomorphic adaptations and is currently known from a mere 11 verified localities spanning the Guizhou, Yunnan, Guangxi, Chongqing, and Hunan provinces. Notably, *Zopherobatrus* is the only known genus of Pselaphinae in China that includes true troglobitic species. Its extreme range restriction, ecological specialization, and complete lack of formal conservation status render it representative of a poorly known group of China's subterranean biodiversity. Quantifying the distributional limits and climate sensitivity of such data-deficient taxa requires spatially explicit predictive tools. While species distribution models (SDMs) have emerged as standard tools for predicting spatial occurrences and guiding conservation planning for rare and endangered species (Elith and Leathwick 2009; Guisan et al. 2013), their application requires careful navigation of conceptual limitations and sensitivity to data quality and species ecology (Araújo and Guisan 2006; Jiménez-Valverde et al. 2008). Despite their proven usefulness in improving the assessment and reporting of rare species (Sousa-Silva et al. 2014), SDM application to strictly subterranean fauna remains in its infancy (Mammola and Leroy 2018). For *Zopherobatrus*, like the majority of cave-adapted invertebrates, distributional limits, climate sensitivity, and

protection status have never been formally assessed, a critical knowledge gap given the rapidly accelerating pace of climate change (Thomas et al. 2004; Thuiller 2004; Bellard et al. 2012) and land-use transformation across the region.

To address these shortfalls, three objectives are pursued here: (1) to model the potential distribution of *Zopherobatrus* under current and future climate scenarios (2050s and 2070s under SSP1-2.6 and SSP5-8.5) using an Ensemble of Small Models (ESMs) framework specifically designed for data-deficient taxa (Breiner et al. 2015, 2018), correcting spatial sampling bias through a target-group background approach based on consubfamilial (Pselaphinae) specimen records (Phillips et al. 2009); (2) to systematically evaluate the effectiveness of China's national and provincial PA network in covering the predicted suitable habitat through a systematic gap analysis; and (3) to identify climate refugia projected to retain habitat suitability across all future scenarios, and assess their current protection status (Keppel et al. 2012). Through this integrated eco-evolutionary assessment, the dual threats of climate change and conservation invisibility faced by subterranean biodiversity are quantified, and evidence-based recommendations for targeted conservation actions are provided.

## Materials and methods

This study follows the ODMAP (Overview, Data, Model, Assessment, Prediction) reporting protocol for species distribution models (Zurell et al. 2020), with the modelling objective classified as "forecast and transfer". An ODMAP checklist describing the workflow used in this study was created at https://odmap.wsl.ch/ and is provided as Suppl. material 1 to improve transparency and reproducibility. Following the framework of Tredennick et al. (2021) and adhering to best-practice guidelines for correlative niche modeling (Sillero et al. 2021; Soley-Guardia et al. 2024), predictive performance was prioritized over parameter inference, with correlative models fitted to current environmental conditions and projected onto future climate scenarios. The underlying ecological hypothesis posits that the distribution of *Zopherobatrus* is primarily constrained by mean annual temperature, precipitation seasonality, aboveground primary productivity, and geographic proximity to karst substrates.

### Species occurrence data

Distribution records of *Zopherobatrus* were compiled from the most recent taxonomic treatment of the genus (Yin 2025) alongside verified museum specimen data. Given the extreme rarity of these taxa and the lack of species-level physiological data, the genus was treated as a single ecological unit, assuming that congeneric species inhabiting similar karst microhabitats share broadly conserved climatic tolerances (Qiao et al. 2017; Smith et al. 2019). After georeferencing and coordinate validation, a total of 11 spatially unique occurrence localities were retained (Fig. 1), spanning Guizhou, Yunnan, Guangxi, Chongqing, and Hunan provinces.

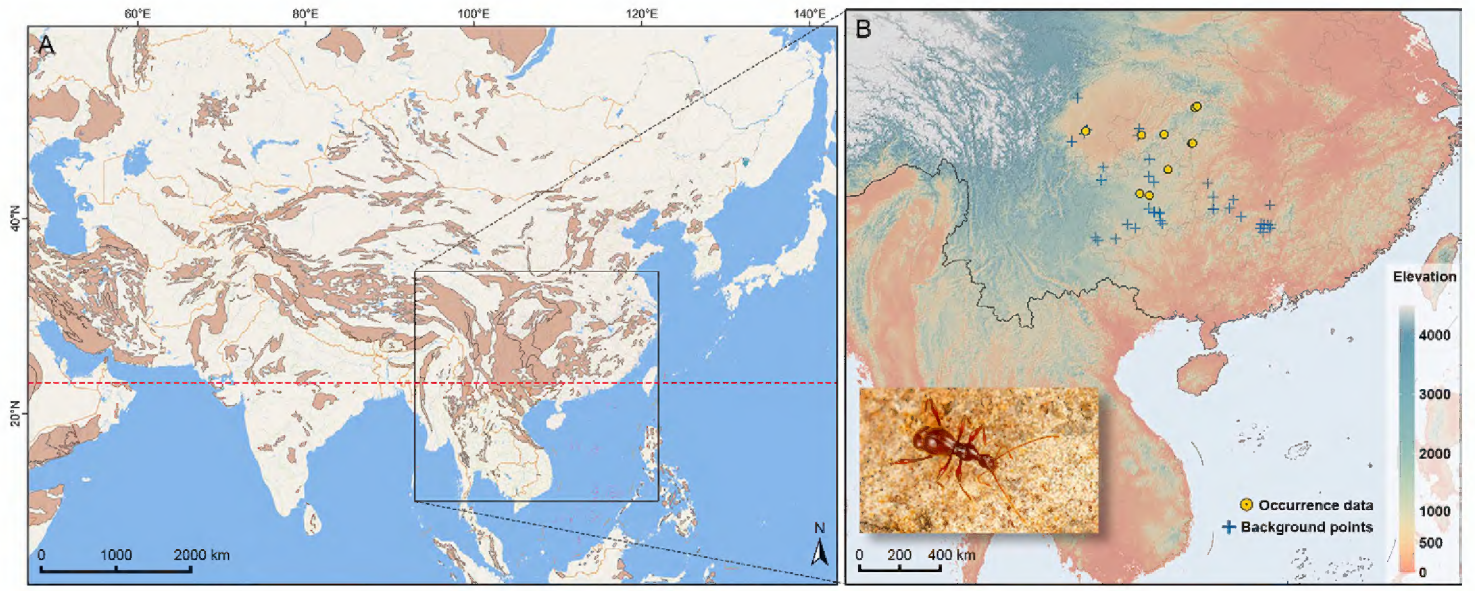


**Figure 1.** Study area and occurrence records of *Zopherobatrus*. **A** regional context showing the distribution of carbonate rock outcrops (brown shading) across Asia, sourced from the World Karst Aquifer Map (WHYMAP; https://www.whymap.org/). The rectangle delineates the study area used for species distribution modeling. The red dashed line indicates the Tropic of Cancer, highlighting the transitional biogeographic position of the genus **B** detailed view of the accessible area (*M* area; 93°–122°E, 10°–35°N), with background coloration representing elevation derived from WorldClim v2.1 30-arc-second digital elevation model. Yellow circles indicate the 11 spatially unique presence records of *Zopherobatrus*. Blue crosses represent target-group background points, used to characterize the available environment while controlling for spatial sampling bias inherent in cave surveys.

To minimize spatial autocorrelation and potential geolocation errors (Graham et al. 2008), occurrence records were thinned to one point per 0.01° grid cell (~1 km) to match the 30 arc-second (~1 km) resolution of the environmental layers, thereby removing potential duplicate records within the same raster pixel. Although formal spatial autocorrelation tests (e.g., Moran's *I*) were not feasible given the extremely small sample size, all 11 localities were situated in distinct county-level administrative units, providing evidence of spatial independence. The extremely small sample size ($N = 11$) poses inherent challenges for correlative species distribution modeling, often leading to overprediction or statistical instability when the number of records falls below recommended minimum thresholds of 14–25 for critical SDM algorithms (Hernandez et al. 2006; Pearson et al. 2007; van Proosdij et al. 2016). Overcoming this paradox requires specialized approaches, such as hierarchical frameworks or ESMs (Lomba et al. 2010; Leroy 2023), which have successfully predicted distributions for highly restricted taxa with minimal occurrences (Breiner et al. 2018).

Background data were generated using the target-group background (TGB) approach, recognizing that pseudo-absence selection strategies may profoundly influence model calibration and predictions (Wisz and Guisan 2009; Barbet-Massin et al. 2012). Georeferenced specimen records of other cavernicolous species of Pselaphinae were used as pseudo-absences, yielding 40 target-group background points. This strategy accounts for spatial bias in sampling effort, as TGB points inherently reflect the

geographic distribution of entomological collecting activity (Phillips et al. 2009). The study area was delimited to 93°–122°E × 10°–35°N, encompassing the known range of *Zopherobatrus* and the broader East Asian karst belt. This extent was chosen to approximate the species' accessible area (*sensu* Barve et al. 2011), thereby reducing the influence of distal environments that the species has never encountered. The modelling approach relies on two critical assumptions: (1) Niche conservatism, assuming that the climatic requirements of *Zopherobatrus* are conserved across time, allowing the model calibrated on current conditions to predict future suitability; and (2) Pseudo-equilibrium, assuming that the observed climatic envelope approximates the genus's realized niche. While dispersal limitations probably constrain its distribution, modeling the realized niche provides the most conservative estimate of climatically suitable habitat.

## Environmental variables

Environmental predictors were selected *a priori* based on their hypothesized ecological relevance to obligate subterranean organisms, following the framework proposed by Mammola and Leroy (2018) and the general principles of ecologically informed variable selection to prevent spurious correlations (Austin 2002; Fourcade et al. 2018; Guisan and Thuiller 2005). Four variables were retained for modeling (Table 1a):

1. Annual Mean Temperature (Bio1): deep cave temperature approximates the local annual mean surface temperature due to thermal equilibrium between surface and subterranean environments (Badino 2010). Bio1 thus serves as a direct proxy for the thermal niche of cave-dwelling ectotherms.

2. Precipitation Seasonality (Bio15): reflects the temporal consistency of groundwater recharge and vadose zone percolation. For stenohygrobic subterranean fauna, high seasonality represents a critical physiological stressor, as prolonged dry intervals can interrupt the essential drip-water supply required to maintain near-saturation humidity levels, regardless of total annual precipitation (Culver and Pipan 2019).

3. Mean Normalized Difference Vegetation Index (NDVI; 5-year average): subterranean ecosystems are allochthonous, relying on surface-derived organic matter as the primary energy source. NDVI serves as a proxy for aboveground primary productivity, which determines the quantity of dissolved organic matter percolating into cave habitats (Simon et al. 2003).

4. Distance to nearest karst outcrop (dist_karst): represents a hard geological constraint, as non-karst lithologies lack the dissolution-formed voids required for subterranean colonization. Distance was calculated from the World Karst Aquifer Map v1.0 (WoKAM; BGR/UNESCO, available at https://www.whymap.org; Chen et al. 2017). Incorporating such distance-to-feature constraints has been shown to reduce overprediction and improve model realism for habitat-specialist taxa (Mendes et al. 2020; Allouche et al. 2008), and combining static geographical constraints with dynamic climatic variables is recommended for climate change studies (Sinclair et al. 2010; Austin and Van Niel 2011; Stanton et al. 2012).

**Table 1.** Environmental variables considered for species distribution modeling of *Zopherobatrus*. Variables were selected *a priori* based on their hypothesized ecological relevance to subterranean organisms (Mammola and Leroy 2018). The highest pairwise Pearson correlation was observed between Bio15 and dist_karst ($|r| = 0.88$); however, all variance inflation factors (VIF) remained below 10 (range: 1.29–7.52), and the ESM framework, which constructs bivariate sub-models, inherently avoids multicollinearity issues within individual models (Breiner et al. 2015). Importance (%) represents Boyce-weighted normalized variable importance derived from ESM sub-model weights (see Methods).

**(a) Environmental variables, selection criteria, and Boyce-weighted importance**

| Variable | Full name | Unit | Source | Resolution | VIF | Importance (%) | Status | Ecological rationale for inclusion/exclusion |
|---|---|---|---|---|---|---|---|---|
| Bio1 | Annual Mean Temperature | °C | WorldClim v2.1 | 30 arc-sec (~1 km) | 1.29 | 15.7 | Retained | Deep cave temperature approximates the local annual mean surface temperature due to thermal equilibrium between epigean and hypogean environments (Mammola et al. 2019). |
| Bio15 | Precipitation Seasonality | CV (%) | WorldClim v2.1 | 30 arc-sec (~1 km) | 7.52 | 26.5 | Retained | Reflects temporal variability in groundwater recharge and subsurface moisture availability. High seasonality implies prolonged dry periods that may desiccate shallow subterranean habitats (Culver and Pipan 2019). |
| NDVI_mean | Mean NDVI (5-year average) | Index (0–1) | MODIS MOD13A2 | 1 km | 1.90 | 31.1 | Retained | Proxy for aboveground primary productivity. Subterranean ecosystems are allochthonous, relying on surface-derived organic matter percolating through the epikarst (Simon et al. 2003). |
| dist_karst | Distance to nearest karst outcrop | m | WoKAM v1.0 | Derived at 30 arc-sec | 6.32 | 26.9 | Retained | Hard geological constraint. Non-karst lithologies lack dissolution-formed voids required for subterranean colonization (Christman et al. 2005). |
| Bio5 | Max Temp. of Warmest Month | °C | WorldClim v2.1 | 30 arc-sec | — | — | Excluded | Cave environments buffer surface temperature extremes; surface maxima do not penetrate the deep subterranean zone (Mammola and Leroy 2018). |
| Bio6 | Min Temp. of Coldest Month | °C | WorldClim v2.1 | 30 arc-sec | — | — | Excluded | Same as Bio5. Surface minimum temperatures are attenuated underground. |
| Bio12 | Annual Precipitation | mm | WorldClim v2.1 | 30 arc-sec | — | — | Excluded | Redundant with Bio15 and NDVI_mean in the study area. Effect captured by Bio15 (temporal distribution) and NDVI_mean (biological integration). |
| Elevation | Altitude above sea level | m | SRTM | 30 arc-sec | — | — | Excluded | Strongly collinear with Bio1 ($r > 0.90$). Temperature is the proximal driver; elevation is a distal correlate (Austin 2002). |

| (a) Environmental variables, selection criteria, and Boyce-weighted importance | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Variable** | **Full name** | **Unit** | **Source** | **Resolution** | **VIF** | **Importance (%)** | **Status** | **Ecological rationale for inclusion/exclusion** |
| Slope | Terrain slope | ° | Derived from DEM | 30 arc-sec | — | — | Excluded | No mechanistic link to subterranean habitat quality. |

| (b) ESM sub-model weights and variable importance decomposition | | | |
|---|---|---|---|
| **Sub-model** | **Variable pair** | **CBI weight** | **Normalized (%)** |
| BIOMOD.1 | Bio1 + Bio15 | 0.041 | 2.3 |
| BIOMOD.2 | Bio1 + NDVI_mean | 0.381 | 21.3 |
| BIOMOD.3 | Bio1 + dist_karst | 0.137 | 7.7 |
| BIOMOD.4 | Bio15 + NDVI_mean | 0.406 | 22.7 |
| BIOMOD.5 | Bio15 + dist_karst | 0.499 | 27.9 |
| BIOMOD.6 | NDVI_mean + dist_karst | 0.323 | 18.1 |

Note: The ESM framework decomposes the full environmental space into six unique pairwise combinations of the four predictor variables, each fitted with a generalized linear model (GLM). Individual sub-model weights were assigned based on their continuous Boyce index (CBI) scores. Normalized variable importance in Panel (a) was calculated as the sum of weights across all sub-models containing a given variable, divided by the total sum of all variable weights. Bio15 and dist_karst, despite their high pairwise correlation ($|r| = 0.88$), are paired in only one sub-model (BIOMOD.5), which received 27.9% of the total ensemble weight.

Climate data (Bio1, Bio15) were obtained from WorldClim v2.1 at 30 arc-second resolution (~1 km; Fick and Hijmans 2017). NDVI data were derived from MODIS MOD13A2 (1 km) (Didan 2021), averaged over five years (1st January 2021 to 31st December 2025) to minimize interannual noise. Pairwise Pearson correlations among the four retained variables were generally low ($|r| \leq 0.50$ for five of six pairs), with the exception of Bio15 and dist_karst ($|r| = 0.88$). Despite this elevated correlation, both variables were retained for ecological and methodological reasons. Ecologically, Bio15 captures temporal moisture variability critical for stenohygrobic cave organisms, whereas dist_karst represents a hard geological constraint on subterranean habitat availability; these two variables reflect markedly different environmental axes and are not substitutable. Because the ESM framework restricts each sub-model to only two predictors, pairwise collinearity between Bio15 and dist_karst does not inflate coefficient variance or bias predictions in the same manner as conventional full-model SDMs. All individual variance inflation factor (VIF) values were below 10 (Dormann et al. 2013).

## Species distribution modeling

Given the extremely small sample size ($N = 11$), an Ensemble of Small Models (ESMs) approach was adopted, which has been demonstrated to outperform standard SDMs and mitigate overfitting for extremely rare species with limited occurrences (Breiner et al. 2015, 2018; Erickson and Smith 2023; Maitner et al. 2026). ESMs overcome the limitations of conventional SDMs for rare species by decomposing the modeling problem into multiple bivariate sub-models, each fitted with only two predictor variables, and then combining them into a weighted ensemble (Pearson et al. 2007). This strategy avoids overfitting by restricting model complexity at the sub-model level while capturing multi-dimensional environmental relationships through aggregation. To assess the robustness of this modelling strategy, an additional MaxEnt models was fitted as a comparative benchmark.

MaxEnt was tuned using ENMeval v2.0 across 48 parameter combinations (four feature-class settings × 12 regularization multiplier values) with jackknife cross-validation, and the optimal model was selected using ΔAICc. Because the sample size was extremely small, ESM-GLM was retained as the primary modelling framework, whereas MaxEnt was used to evaluate whether the main ecological conclusions were sensitive to algorithm choice.

All ESMs were fitted using generalized linear models (GLMs; Nelder and Wedderburn 1972) as the base algorithm, specifying a binomial error distribution and a logit link function. To balance model complexity with the limited sample size, bivariate sub-models were restricted to linear and quadratic terms without interactions, favoring response curve simplicity to improve transferability and ecological interpretability (Merow et al. 2014). The models were implemented in the 'ecospat' package (v4.1.3; Di Cola et al. 2017) within the 'biomod2' framework (v4.3-4-5; Thuiller et al. 2009; Guéguen et al. 2026) in R v4.5.2 (R Core Team 2025). For each bivariate sub-model, 70% of the data were used for calibration and 30% for evaluation (Fielding and Bell 1997; Franklin 2010a), repeated across 10 cross-validation runs ('NbRunEval = 10'). Individual sub-models were weighted by their continuous Boyce index (CBI; Hirzel et al. 2006) scores, and a Boyce-weighted ensemble prediction was generated to reduce inter-model uncertainty (Araújo and New 2007; Hao et al. 2019).

## Model evaluation

Model performance was assessed using the continuous Boyce index (CBI), which is specifically recommended for presence-only models as it does not require true absence data (Hirzel et al. 2006) and avoids the inflated performance estimates associated with AUC (area under the receiver operating characteristic curve) in SDMs (Lobo et al. 2008). In addition, CBI is preferred over threshold-dependent metrics like TSS, which are sensitive to prevalence (Swets 1988; Allouche et al. 2006). A null-model significance test was conducted: 999 null models were generated by randomly permuting occurrence localities within the study area, and the empirical CBI was compared against this null distribution to obtain a *P*-value (Raes and ter Steege 2007). The 10$^{th}$-percentile training presence threshold (*P10* = 608 on a 0–1000 suitability scale) was applied to convert continuous suitability predictions into binary presence/absence maps (Liu et al. 2013). This threshold was selected to define 'meaningful suitability' conservatively, ensuring that the top 90% of known occurrence localities are included in the protection envelope while allowing for a 10% error margin potentially caused by sink populations or georeferencing inaccuracies. Additionally, an ecological plausibility check (Zurell et al. 2020) was performed by visually inspecting the response curves of the final ensemble model to ensure that the modeled relationships (e.g., decreasing suitability with increasing distance from karst) were consistent with biological knowledge of the genus.

To evaluate the robustness of model selection, MaxEnt (Phillips et al. 2006, 2017) was additionally tested with exhaustive hyperparameter optimization using ENMeval v2.0 (Kass et al. 2021). A total of 48 parameter combinations were evaluated: four feature class configurations (L, LQ, H, LQH) crossed with 12 regularization multiplier values

(0.5–6.0, step 0.5; Merow et al. 2013), using jackknife (leave-one-out, $N = 11$) cross-validation. Model selection based on the small-sample corrected Akaike Information Criterion ($\Delta$AICc = 0; Warren and Seifert 2011) identified a Linear + Quadratic model (RM = 1.0, 3 non-zero coefficients) as optimal. This model achieved a Continuous Boyce Index (CBI) of 0.969 (vs. 0.910 for the ESM-GLM), confirming that both approaches detected a strong, genuine ecological signal in the occurrence data. However, the optimal MaxEnt model exhibited an omission rate of 27.3% (OR10 = 0.273) at the 10$^{th}$ percentile training threshold, which was considerably higher than the expected 10%, suggesting spatial prediction instability typical of small-sample MaxEnt applications (Moreno-Amat et al. 2015; Qiao et al. 2015; Morales et al. 2017). Given that the ESM-GLM framework provides (1) explicit variable importance decomposition through ensemble-weighted bivariate sub-models, (2) avoidance of multicollinearity by design, and (3) a theoretical foundation specifically developed for rare species with few occurrence records (Breiner et al. 2015, 2018), ESM-GLM was retained as the primary modelling framework. The concordance of both approaches (CBI > 0.9) strengthens confidence that the habitat suitability predictions are robust and not artefacts of algorithmic choice (Suppl. materials 2, 4).

## Future projections

Recognizing that the choice of future climate scenarios significantly influences predictive modeling (Beaumont et al. 2008), current habitat suitability was projected onto four future climate scenarios representing two time periods (2041–2060 ["2050s"] and 2061–2080 ["2070s"]) under two Shared Socioeconomic Pathways: SSP1-2.6 (low emissions, sustainability) and SSP5-8.5 (high emissions, fossil-fueled development) (O'Neill et al. 2017; Riahi et al. 2017). Future bioclimatic variables were obtained from the WorldClim v2.1 CMIP6 downscaled climate projections (https://www.worldclim.org/data/cmip6/cmip6climate.html) at 30 arc-second resolution, based on three general circulation models (GCMs; Eyring et al. 2016): BCC-CSM2-MR (Wu et al. 2019), MIROC6 (Tatebe et al. 2019), and MRI-ESM2-0 (Yukimoto et al. 2019), and ensemble-averaged to reduce inter-model uncertainty (Araújo and New 2007). Only Bio1 and Bio15 were updated for future scenarios; NDVI and dist_karst were held constant under the conservative assumption that geological substrates remain unchanged and that vegetation productivity changes are not reliably projected at the required resolution. This assumption is explicitly acknowledged as a limitation.

To assess the reliability of model extrapolation into novel climatic conditions, multivariate environmental similarity surface (MESS) analysis was performed following Elith et al. (2010), acknowledging that rigorous assessment of extrapolation risks is critical for climate change projections (Velazco et al. 2024). Areas with negative MESS values, indicating environmental conditions outside the range of the training data, were flagged as zones of model extrapolation uncertainty.

For each scenario, the total area of suitable habitat (km$^2$) was calculated using latitude-corrected cell areas. Range centroid shifts were computed as the Haversine distance between the weighted centroids of current and future binary distribution maps.

Spatial turnover was decomposed into three components: stable habitat (currently and future suitable), lost habitat (currently suitable, future unsuitable), and gained habitat (currently unsuitable, future suitable).

### Conservation gap analysis

The effectiveness of existing protected areas (PAs) in covering *Zopherobatrus* habitat was evaluated through a systematic gap analysis. The boundaries of China's national and provincial nature reserves were obtained from a georeferenced dataset (China Nature Reserve Specimen Resource Sharing Platform 2024; Zenodo DOI: 10.5281/zenodo.14875797). PA polygons were rasterized to 30 arc-second resolution and overlaid with the binary suitability map to calculate: (1) the total area and proportion of suitable habitat within PAs; and (2) the number of known occurrence localities falling within PA boundaries.

Climate refugia were defined as areas predicted to remain suitable (i.e., above the *P10* threshold) under all four future scenarios (2 SSPs × 2 time periods; Keppel et al. 2012). The proportion of climate refugia covered by existing PAs was then calculated. Conservation priority was assigned on a five-level scale: (1) unsuitable—below *P10* threshold; (2) low priority—suitable under current conditions but outside climate refugia; (3) moderate priority—suitable within climate refugia but unprotected; (4) high priority—high suitability within refugia and within existing protected areas; and (5) critical conservation gap—high suitability within refugia but outside any protected area.

All spatial analyses were conducted in R v4.5.2 using the 'terra' (v1.8-93; Hijmans et al. 2026) and the 'sf' (v1.0-24; Pebesma 2018; Pebesma and Bivand 2023) packages.

### Data Resources

All occurrence records, environmental layers, and the full R code workflow required to reproduce the ESM and gap analysis have been archived in Zenodo at https://doi.org/10.5281/zenodo.18737001.

## Results

### Model performance and variable contributions

The ESM ensemble model demonstrated robust predictive performance. The continuous Boyce index (CBI) was 0.910, indicating a strong positive correlation between model predictions and the observed distribution of *Zopherobatrus*. The null-model significance test confirmed that this performance was significantly better than expected by chance ($P = 0.001$, based on 999 null permutations; Suppl. materials 3: fig. S2a, b, 5). MESS analysis revealed that all 11 occurrence localities fell within the interpolation range of the training data (MESS ≥ 0), supporting the reliability of model predictions within the known distribution area (Suppl. material 3: fig. S2c).

Among the four predictor variables, NDVI_mean and distance to the nearest karst outcrop (dist_karst) showed the highest Boyce-weighted normalized variable importance values (31.1% and 26.9%, respectively; Table 1) and were therefore the strongest contributors to predicted habitat suitability, reflecting the dependence of *Zopherobatrus* on carbonate substrates and its sensitivity to surface-derived organic matter inputs. Precipitation seasonality (Bio15) and annual mean temperature (Bio1) also contributed substantially to the ensemble predictions (26.5% and 15.7%, respectively; Table 1), consistent with the hypothesis that subterranean organisms are indirectly influenced by surface moisture regimes and thermal conditions.

## Current distribution and future projections

Under current climatic conditions, the ESM ensemble predicted 226,475 km$^2$ of climatically suitable habitat for *Zopherobatrus* (Fig. 2). The predicted suitable area was concentrated across the karst regions of southwest China, spanning Guizhou, Yunnan, Guangxi, western Hunan, Chongqing, and parts of southern Sichuan, with the highest suitability values clustered along the deeply incised montane regions of the Guizhou–Hunan–Chongqing border area.

Under all four future climate scenarios, substantial habitat contraction was projected (Table 2, Fig. 3). Under the optimistic pathway (SSP1-2.6), suitable habitat was projected to decline by 39.0% by the 2050s (138,152 km$^2$) and by 36.4% by the 2070s

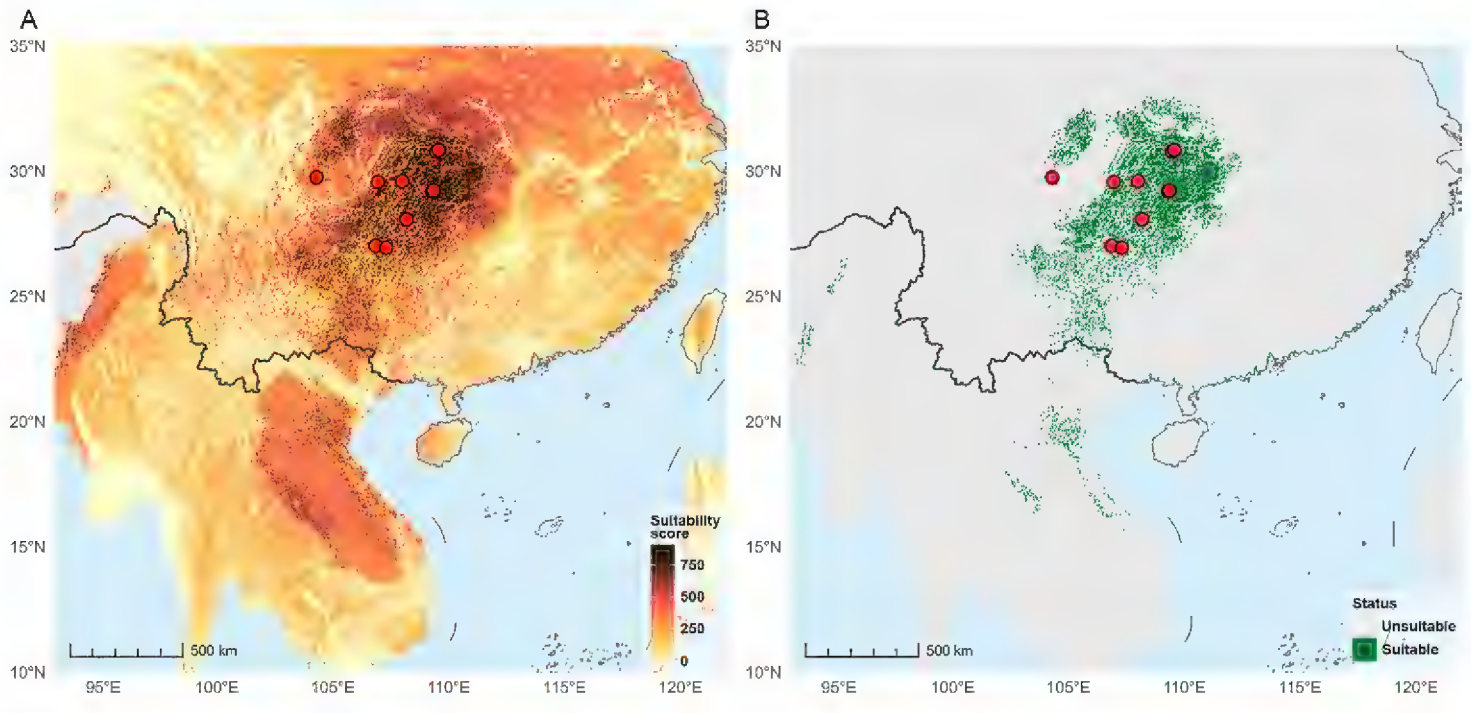


**Figure 2.** Current habitat suitability for *Zopherobatrus* predicted by the Ensemble of Small Models (ESMs) framework. **A** continuous suitability scores (0–1,000 scale), with higher values (darker shading) indicating greater environmental suitability **B** binary habitat classification based on the 10$^{th}$-percentile training presence threshold (*P10* = 608), delineating 226,475 km$^2$ of climatically suitable habitat (green). Red circles denote the 11 spatially unique occurrence localities. The national boundary of China is shown for reference.

**Table 2.** Projected changes in climatically suitable area for *Zopherobatrus* under four future climate scenarios. Predictions were generated using an ensemble of three general circulation models (BCC-CSM2-MR, MIROC6, MRI-ESM2-0) under two Shared Socioeconomic Pathways (SSP1-2.6 and SSP5-8.5) for the 2050s (2041–2060) and 2070s (2061–2080). Suitable area was defined using the 10$^{th}$-percentile training presence threshold (*P10* = 608).

| Scenario | Suitable area (km²) | Change (km²) | Change (%) | Stable (km²) | Retained (%) | Lost (km²) | Gained† (km²) | Centroid shift (km) |
|---|---|---|---|---|---|---|---|---|
| Current | 226,475 | — | — | — | — | — | — | — |
| SSP1-2.6 (2050s) | 138,152 | -88,324 | -39.0 | 106,546 | 47.0 | 119,930 | 31,606 | 69.7 |
| SSP1-2.6 (2070s) | 143,960 | -82,515 | -36.4 | 109,240 | 48.2 | 117,236 | 34,720 | 52.4 |
| SSP5-8.5 (2050s) | 113,351 | -113,125 | -50.0 | 77,534 | 34.2 | 148,942 | 35,817 | 92.1 |
| SSP5-8.5 (2070s) | 83,528 | -142,948 | -63.1 | 54,286 | 24.0 | 172,190 | 29,242 | 141.1 |

†Gained habitat is ecologically unrealizable for *Zopherobatrus*. The retained habitat ratio (Retained %) therefore represents the primary metric of future habitat viability.

(143,960 km²). This counterintuitive partial recovery reflects the emission stabilization inherent to SSP1-2.6, under which radiative forcing peaks around mid-century and begins to decline, allowing a subset of previously marginal habitats to re-enter the suitability envelope. Under the high-emission pathway (SSP5-8.5), more severe contractions were projected, i.e., a 50.0% reduction by the 2050s (113,351 km²) and a 63.1% reduction by the 2070s (83,528 km²).

The range centroid was predicted to shift northwestward toward higher elevations under all scenarios, with the magnitude of displacement ranging from 52 km (SSP1-2.6, 2070s) to 141 km (SSP5-8.5, 2070s). Spatial turnover analysis revealed that habitat loss was concentrated along the southern and eastern margins of the current distribution (Fig. 4). Although newly suitable areas (hereafter 'gained' habitat) were projected to appear along the northern periphery, these areas are probably ecologically irrelevant because *Zopherobatrus* comprises obligate cave-dwelling organisms with negligible dispersal capacity and strict dependence on karst substrates; colonization of disjunct newly suitable areas is functionally impossible (Christman et al. 2005). Consequently, a more biologically meaningful metric for this genus is the 'retained habitat' ratio, representing the proportion of currently suitable habitat that remains suitable under future conditions. This ratio ranged from as low as 24.0% (SSP5-8.5, 2070s) to 48.2% (SSP1-2.6, 2070s), indicating that over half to three-quarters of the current suitable habitat will be lost without compensatory range expansion (Table 2).

## Conservation gap analysis

The gap analysis revealed a pronounced mismatch between the distribution of suitable habitat and the coverage of existing protected areas (Table 3a). Of the 226,475 km² of currently suitable habitat, only 5,841 km² (2.6%) fell within the boundaries of national or provincial nature reserves. More critically, none of the 11 known occurrence localities of *Zopherobatrus* was located within any protected area (Table 3b).

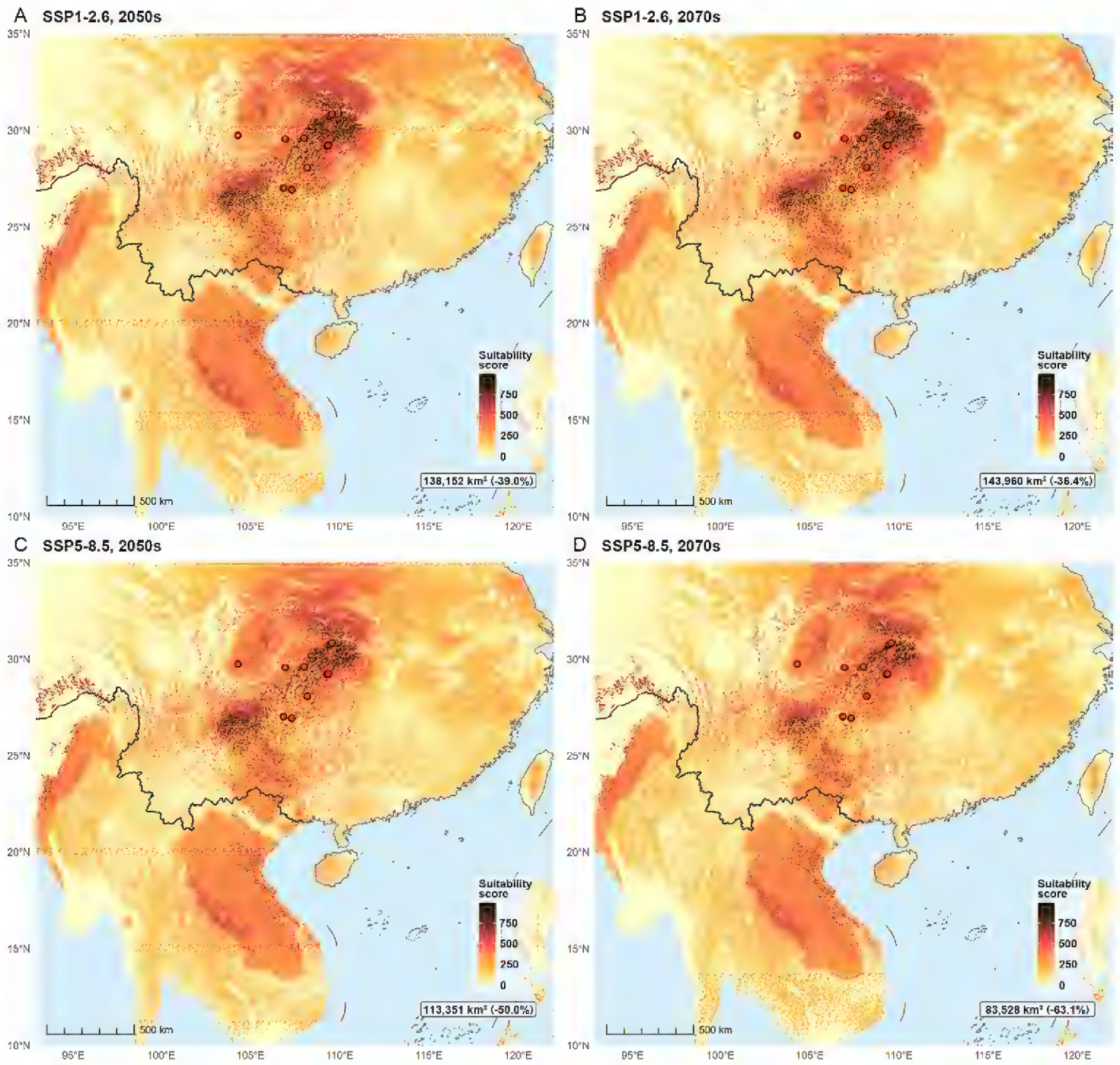


**Figure 3.** Projected habitat suitability for *Zopherobatrus* under four future climate scenarios. **A** SSP1-2.6, 2050s **B** SSP1-2.6, 2070s **C** SSP5-8.5, 2050s **D** SSP5-8.5, 2070s. Suitability values (0–1,000 scale) represent ESM ensemble predictions based on the mean of three CMIP6 GCMs (MIROC6, MRI-ESM2-0, BCC-CSM2-MR). NDVI and distance to karst were held constant across all scenarios. The projected suitable area and percentage change relative to the current distribution (226,475 km²) are annotated in each panel. Red circles denote known occurrence localities.

A total of 49,299 km² of climate refugia, i.e., areas predicted to remain suitable under all four future scenarios, were identified (Fig. 5). These refugia were concentrated in the deeply incised karst valleys of the Guizhou–Hunan–Chongqing border area and parts of northern Guangxi. However, 95.9% of these climate refugia (47,299 km²) were found to lie outside any existing protected area. Within these unprotected refugia, 12,611 km² of high-suitability habitat was identified as critical conservation gaps (Table 3a). Among the 11 known localities, only one was situated within a climate refugium. These critical gaps, characterized by high suitability, refugial status, and absence of formal protection, represent the highest-priority targets for future conservation intervention.

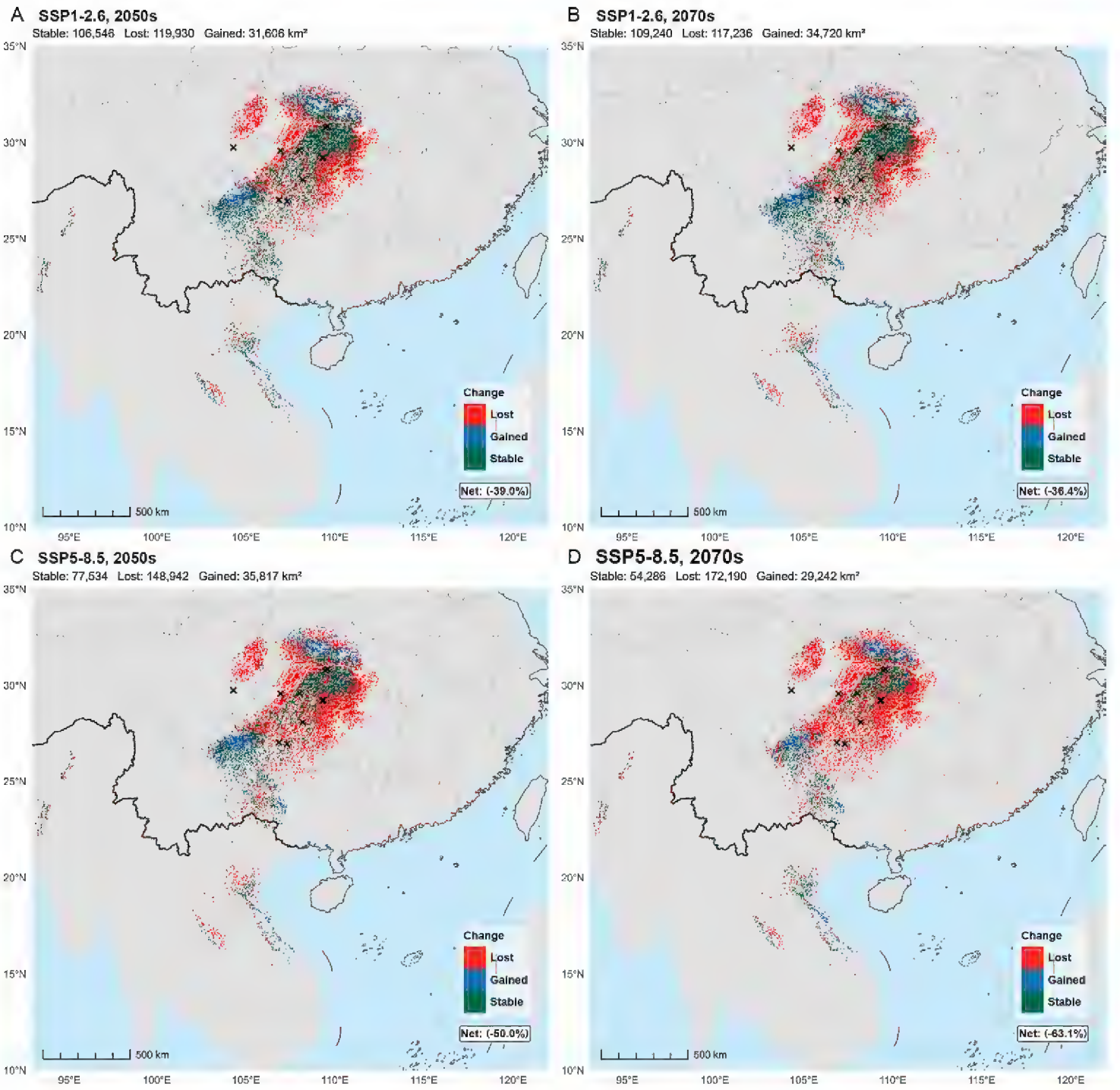


**Figure 4.** Spatial turnover in *Zopherobatrus* suitable habitat under four future climate scenarios. **A** SSP1-2.6, 2050s **B** SSP1-2.6, 2070s **C** SSP5-8.5, 2050s **D** SSP5-8.5, 2070s. Pixels are classified as stable (green; suitable under both current and future conditions), lost (red; currently suitable but projected to become unsuitable), or gained (blue; currently unsuitable but projected to become suitable). The area ($km^2$) of each category is shown above each panel, with the net percentage change annotated in the lower right. Black crosses indicate known occurrence localities.

## Discussion

### Convergence of climate vulnerability and conservation neglect

The results presented here reveal a deeply concerning convergence of threats faced by *Zopherobatrus* and, by extension, the broader cave-obligate fauna of China's karst landscapes. On one axis, the bioclimatic envelope is projected to shift up to 141 km northwestward. For *Zopherobatrus*, a genus including blind, flightless beetles restricted to iso-

**Table 3.** Conservation gap analysis for *Zopherobatrus* suitable habitat and climate refugia. Protected area coverage was calculated using national and provincial nature reserve boundaries (Zenodo DOI: https://doi.org/10.5281/zenodo.14875796). Climate refugia were defined as areas predicted to remain suitable under all four future scenarios (SSP1-2.6 and SSP5-8.5 for the 2050s and 2070s). Suitable habitat was delineated using the 10th-percentile training presence threshold (*P10* = 608). High suitability was defined as scores above the 75th percentile of values at known occurrence localities.

**(a) Habitat area and protection coverage**

| Category | Area (km²) | Proportion | Reference |
|---|---|---|---|
| Total suitable habitat (current) | 226,475 | 100% | — |
| High suitability (core habitat) | 35,797 | 15.8% | of total suitable |
| Moderate suitability | 190,678 | 84.2% | of total suitable |
| Climate refugia | 49,299 | 21.8% | of total suitable |
| Protected suitable habitat | 5,841 | 2.6% | of total suitable |
| Protected climate refugia | 2,000 | 4.1% | of climate refugia |
| Critical conservation gap (unprotected core refugia) | 12,611 | 25.6% | of climate refugia |
| Occurrence localities in refugia | 1 / 11 | 9.1% | of known localities |
| Occurrence localities in protected areas | 0 / 11 | 0% | of known localities |

**(b) Protection and refugium status of the 11 known occurrence localities**

| Locality | Latitude, Longitude | Suitability | Refugium | Priority |
|---|---|---|---|---|
| 1 | 29.5762°N, 106.9478°E | 650 | No | Low |
| 2 | 27.0236°N, 106.8547°E | 608 | No | Low |
| 3 | 29.7491°N, 104.3027°E | 362† | No | Unsuitable |
| 4 | 30.7833°N, 109.4695°E | 841 | No | Low |
| 5 | 30.8450°N, 109.5725°E | 844 | No | Low |
| 6 | 28.0789°N, 108.1881°E | 859 | No | Low |
| 7 | 26.9461°N, 107.3131°E | 790 | Yes | Moderate |
| 8 | 29.2088°N, 109.3061°E | 822 | No | Low |
| 9 | 29.2076°N, 109.3043°E | 788 | No | Low |
| 10 | 29.2272°N, 109.3584°E | 806 | No | Low |
| 11 | 29.6031°N, 108.0031°E | 769 | No | Low |

†Below *P10* threshold (608); consistent with the expected 10% omission rate for an 11-point dataset. Conservation priority classes correspond to Fig. 5: Unsuitable = below *P10* threshold; Low = suitable, non-refugium; Moderate = moderate suitability within refugium. None of the 11 localities falls within an existing protected area.

lated karst towers, such dispersal is functionally impossible. Unlike surface organisms that might track shifting isotherms (Thuiller 2004), these troglobionts are geologically trapped; their inability to cross non-karstic barriers transforms this 'range shift' into a substantial extinction debt (Christman et al. 2005; Mammola and Leroy 2018). On the other axis, the existing PA network provides virtually no coverage: only 2.6% of currently suitable habitat is formally protected, and not a single known locality falls within a reserve. Notably, this level of protection is even lower than the already modest global overlap between protected areas and subterranean ecosystems (6.9%; Sánchez-Fernández et al. 2021), pointing to the severity of the mismatch documented for *Zopherobatrus*. Collectively, this "double jeopardy"—simultaneous climate-driven habitat loss and near-total absence of legal protection—mirrors patterns documented for other subterranean taxa globally (Mammola et al. 2019; Vaccarelli et al. 2023) but is here quantified for the first time for a cave-obligate beetle group in one of the world's most species-rich karst regions.

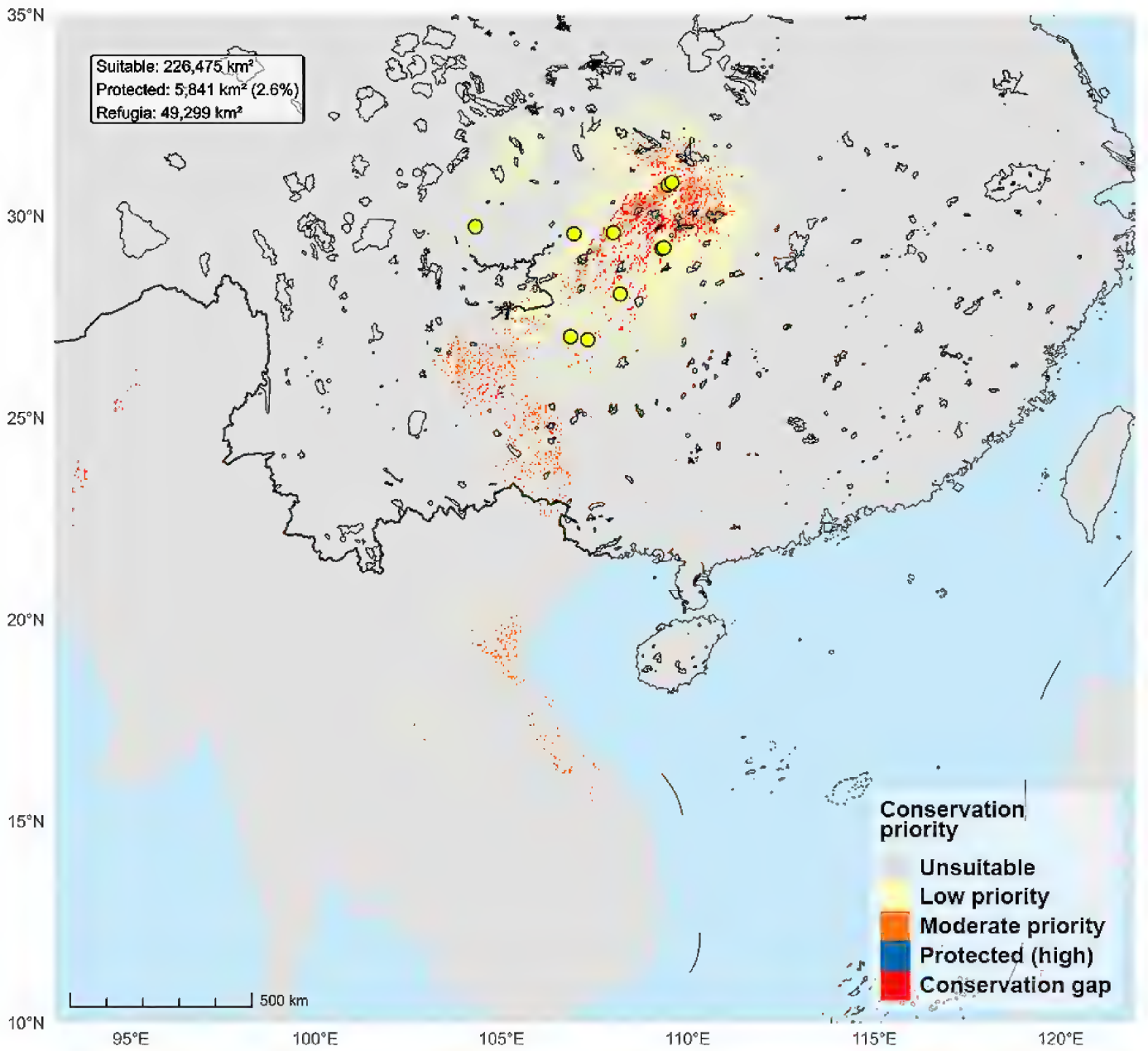


**Figure 5.** Conservation priority map for *Zopherobatrus*. Five priority levels are displayed: (1) unsuitable habitat (grey); (2) low priority—suitable under current conditions but outside climate refugia (yellow); (3) moderate priority—suitable within climate refugia but unprotected (orange); (4) protected (high)—high suitability within refugia and within existing protected areas (blue); (5) conservation gap—high suitability within refugia but outside any protected area, representing the most critical targets for future conservation action (red). Summary statistics are annotated: total suitable area = 226,475 km$^2$; protected suitable area = 5,841 km$^2$ (2.6%); climate refugia = 49,299 km$^2$. Protected area boundaries are overlaid in black. Yellow circles denote known occurrence localities.

As argued in the Results, the retained habitat ratio (24.0–48.2%) provides a more meaningful assessment than the net change metric, which implicitly credits biologically impossible colonization of disjunct 'gained' habitat. This distinction suggests that area-based targets should be based on retained (stable) habitat alone.

The extremely low PA coverage identified in this study cannot be attributed to insufficient PA density in southwest China, which in fact hosts a relatively high concentration of nature reserves. Rather, it reflects the fundamental mismatch between PA placement criteria, which have historically prioritized surface vegetation, charismatic vertebrates, and aboveground ecosystem services, and the spatial distribution of subterranean biodiversity, which is governed by subsurface geological and

hydrogeological features, highlighting a fundamental vertical disconnect in current conservation strategies (Mammola et al. 2019, 2020, 2024). This finding points to the need for conservation frameworks that explicitly incorporate belowground biodiversity as a criterion in PA network design.

## Climate refugia and their conservation significance

The identification of 49,299 km$^2$ of climate refugia, areas projected to retain climatic suitability across all four future scenarios, provides a spatially explicit foundation for conservation prioritization (Keppel et al. 2012). These refugia are concentrated along the montane regions of the Guizhou–Hunan–Chongqing border and parts of northern Guangxi, characterized by complex topography, high elevation gradients, and extensive karst development. The concentration of refugia in topographically heterogeneous landscapes is consistent with the "terrain ruggedness hypothesis" (Dobrowski 2011; Morelli et al. 2020), which posits that complex terrain provides micro-climatic buffering against regional warming by offering a mosaic of thermal conditions over short distances.

However, the finding that 95.9% of these refugia lie outside any existing PA warrants attention. If climate refugia represent the last remaining strongholds for cave-adapted species under future warming, the current absence of protection renders these areas particularly vulnerable to additional anthropogenic pressures, including quarrying of limestone for the cement industry, groundwater extraction, and agricultural encroachment—all of which are prevalent across the karst landscapes of southwest China (Jiang et al. 2014). The conservation value of refugia can only be realized if they are translated into actionable spatial plans.

## Conservation implications and recommendations

The findings of this study provide the quantitative baseline required for evidence-based conservation of subterranean ecosystems (Mammola et al. 2022), and three actionable conservation recommendations are here proposed. First, the establishment of targeted micro-reserves (*sensu* Lumaret and Lobo 1996) is recommended in the core areas of identified climate refugia, particularly in the deeply incised karst valleys of the Guizhou–Hunan–Chongqing border region. Unlike conventional large-scale PAs, micro-reserves are small, precisely delineated conservation units designed to protect specific habitats or populations of narrow-range endemics. For cave-dwelling taxa, such reserves would need to include not only cave entrances but also the surrounding surface catchment area, given the dependence of subterranean ecosystems on surface-derived energy and hydrological inputs, a "vertical conservation unit" approach that integrates above- and belowground biodiversity protection (Culver and Pipan 2019). It should be noted that even within existing protected areas, the legal framework governing subterranean resource use in China remains ambiguous; surface-level nature reserve designations do not necessarily confer explicit protection

against subsurface mining, quarrying, or groundwater extraction, potentially rendering the 2.6% coverage estimate an overestimate of effective protection. Second, the inclusion of *Zopherobatrus* in China's List of Key Protected Wild Animals (《国家重点保护野生动物名录》) is recommended. None of the known *Zopherobatrus* species currently holds any formal conservation status at the national or provincial level. Given that it represents the only evolutionary lineage of true troglobitic Pselaphinae in China, alongside the genus's extreme range restriction, ecological specialization, and high vulnerability to climate change, formal legal protection would provide a regulatory framework for habitat preservation and would elevate the visibility of cave-dwelling invertebrates in China's conservation discourse. Third, the systematic integration of subterranean fauna into regional conservation prioritization frameworks is urged. The distributional data and suitability maps generated in this study provide a template for incorporating cave-adapted invertebrates into conservation planning. By overlaying suitability maps for multiple cave-obligate taxa, including other troglobitic beetles, spiders, and crustaceans, future studies could delineate "subterranean biodiversity hotspots" and identify areas where conservation investments would yield the greatest returns for belowground biodiversity as a whole.

## Future directions for subterranean conservation

To effectively conserve subterranean biodiversity in the face of rapid environmental change, the scientific community must bridge the gap between discovery and protection, a challenge often described as the "taxonomic impediment" (Cardoso et al. 2011). Although SDM methodologies have not yet reached widespread popularity among cave biologists and subterranean ecologists, recent reviews of methodological developments and ongoing challenges (Franklin 2010b, 2013; Fois et al. 2018) confirm that SDMs remain effective for guiding targeted field surveys of poorly known species at local scales. A "discovery-description-protection" workflow for cave insects is proposed:

1. Accelerated discovery: strengthening collaboration between professional biologists and speleological societies is crucial. "Citizen scientists" (e.g., cavers) play a pivotal role in accessing deep cave zones that are logistically challenging for researchers (Mammola et al. 2019).
2. Integrative description: the rate of species description must keep pace with discovery. An integrative taxonomic approach combining classical morphology with molecular data (e.g., *COI* markers) is advocated to resolve cryptic diversity and accelerate the description of new taxa (Mammola et al. 2020).
3. Precautionary protection: conservation assessments should not wait for perfect data. The use of SDM-derived range estimates to conduct preliminary threatened species list assessments for data-deficient invertebrates is recommended (e.g., IUCN Red List; Bachman et al. 2011), ensuring that potential extinction risks are flagged early in the policy-making process.

## Limitations

Although the ESM framework precludes within-model collinearity by construction, the high correlation between Bio15 and dist_karst ($|r|$ = 0.88; Suppl. material 3: fig. S2d) warrants acknowledgment. Boyce-weighted variable importance scores (Table 1b) showed normalized contributions of NDVI_mean (31.1%), dist_karst (26.9%), Bio15 (26.5%), and Bio1 (15.7%). The combined importance of Bio15 and dist_karst (53.4%) exceeds that of other variable pairs, and readers should interpret this with appropriate caution. Nevertheless, these two variables reflect fundamentally different ecological axes, i.e., dist_karst captures a hard geological constraint on subterranean habitat existence, whereas Bio15 captures temporal moisture variability critical for stenohygrobic cave organisms; and their response curves exhibit qualitatively different shapes (monotonically decreasing vs. unimodal), indicating ecological complementarity rather than pure redundancy.

The results must also be interpreted within the context of several inherent limitations characteristic of modeling rare, subterranean taxa. First, the small sample size presents a challenge for statistical inference, and species' ecological traits may further modulate model accuracy (McPherson and Jetz 2007; Moudrý et al. 2024). Reliance on presence-only data limits the ability to strictly quantify commission errors (Elith et al. 2020). While the ESM framework considerably improves model performance for rare species (Breiner et al. 2015, 2018), the resulting models describe the realized niche, which may be narrower than the fundamental niche given the dispersal limitations inherent to karst-restricted troglobites (Pearson et al. 2007). Notably, one training locality (Locality 3, Sichuan) scored below the *P10* threshold (362 vs. 608), consistent with the expected 10% omission rate for an 11-point dataset. Its outlier environmental profile suggests a marginal population persisting in a microrefugium, highlighting the resolution limitations of correlative SDMs for subterranean organisms (Mammola and Leroy 2018).

Second, the use of surface climate variables to model subterranean niches assumes surface-subsurface thermal coupling. While empirical studies confirm such coupling in shallow cave zones (Badino 2010; Mammola et al. 2018), deep subterranean environments involve significant thermal inertia with lag times of decades to centuries (Badino 2010). The projections therefore represent an 'equilibrium endpoint' rather than a prediction of imminent habitat loss. This thermal buffering defines a critical intervention window during which proactive protection measures could be implemented before irreversible thresholds are breached. Third, the "Linnean" and "Wallacean" shortfalls (Cardoso et al. 2011; Hortal et al. 2015; Ficetola et al. 2019) mean the models likely underestimate the true distribution as additional populations/undescribed species most probably exist in unexplored karst systems.

Finally, by holding NDVI constant, the models implicitly assume that the allochthonous organic matter flux sustaining subterranean food webs (Simon et al. 2003) remains unchanged under future climate scenarios, yielding conservative estimates of habitat loss. Climate-driven aridification across southwest China's karst landscapes

(Jiang et al. 2014) would directly diminish cave energy input, potentially driving a 'starvation before overheating' pathway. To quantify this anchoring effect, a supplementary ESM excluding NDVI (Bio1 + Bio15 + dist_karst; Suppl. material 6) was constructed. The No-NDVI ESM achieved comparable current-range discrimination (CBI = 0.916 vs. 0.910), but retained habitat ratios declined by 14.3–16.2 percentage points across all scenarios (9.6–33.9% vs. 24.0–48.2%), confirming that the main model's retention estimates represent an upper bound. Under SSP5-8.5 by the 2070s, only 9.6% habitat retention was projected, suggesting near-complete range collapse when the assumption of constant surface productivity is relaxed. The full 4-variable model is retained as the primary framework because the ecological rationale for including NDVI is well-established for subterranean organisms (Simon et al. 2003; Mammola and Leroy 2018).

Future modeling efforts should incorporate dynamic vegetation-subsurface energy coupling to more fully characterize extinction risk for cave-obligate fauna. Collectively, these limitations imply that the suitability maps should be viewed as testable hypotheses identifying candidate refugia for future field verification, rather than definitive predictions of species persistence (Guillera-Arroita et al. 2015; Jarnevich et al. 2015; Lee-Yaw et al. 2022; Velazco et al. 2024).

## Conclusion

This study provides the first quantitative assessment of climate vulnerability and conservation coverage for cave-obligate beetles in the karst biodiversity hotspot of southwest China. Through the application of an Ensemble of Small Models framework to the rare and data-deficient genus *Zopherobatrus*, three key findings have been established: (1) only 24.0–48.2% of climatically suitable habitat was projected to be retained under future climate scenarios (retained habitat ratio), an upper-bound estimate reflecting the dampening effect of the temporally static NDVI predictor. A supplementary sensitivity ESM excluding NDVI confirmed that pure climatic and geological signals drive even greater habitat loss (9.6–33.9% retained; Suppl. material 6), with NDVI anchoring accounting for 14–16 percentage points of apparent habitat stability. Notably, the conventional net area contraction metric (–36.4% to –63.1%) further masks functional habitat loss because climatically "gained" areas are ecologically inaccessible to this group; the range centroid shifted up to 141 km northwestward; (2) the existing protected area network covers only 2.6% of suitable habitat, with none of the 11 known localities formally protected; and (3) 95.9% of identified climate refugia lie outside any existing reserve. These results quantify a "double jeopardy" for subterranean biodiversity—the simultaneous erosion of suitable habitat by climate change and the near-total absence of legal protection. The establishment of micro-reserves centered on climate refugia, the formal listing of all described species of *Zopherobatrus* as protected taxa, and the systematic incorporation of subterranean fauna into conservation planning are urgently recommended to address this critical blind spot in regional biodiversity protection.

## Acknowledgements

I thank Stefano Mammola (Water Research Institute, National Research Council, Verbania, Italy) for thoughtful comments on the draft manuscript, which improved the quality of the paper. The present study was supported by grants from the National Natural Science Foundation of China (Nos. 32370465, 31812965). Gemini 3.1 Pro was used to polish the English narrative and improve readability; however, the author retains full responsibility for the scientific facts, accuracy, and overall integrity of the reported work.

## References

Allouche O, Steinitz O, Rotem D, Rosenfeld A, Kadmon R (2008) Incorporating distance constraints into species distribution models. Journal of Applied Ecology 45(2): 599–609. https://doi.org/10.1111/j.1365-2664.2007.01445.x

Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). Journal of Applied Ecology 43(6): 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Araújo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. Journal of Biogeography 33(10): 1677–1688. https://doi.org/10.1111/j.1365-2699.2006.01584.x

Araújo MB, New M (2007) Ensemble forecasting of species distributions. Trends in Ecology & Evolution 22(1): 42–47. https://doi.org/10.1016/j.tree.2006.09.010

Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological Modelling 157(2–3): 101–118. https://doi.org/10.1016/S0304-3800(02)00205-3

Austin MP, Van Niel KP (2011) Improving species distribution models for climate change studies: variable selection and scale. Journal of Biogeography 38(1): 1–8. https://doi.org/10.1111/j.1365-2699.2010.02416.x

Bachman S, Moat J, Hill AW, de la Torre J, Scott B (2011) Supporting Red List threat assessments with GeoCAT: geospatial conservation assessment tool. ZooKeys 150: 117–126. https://doi.org/10.3897/zookeys.150.2109

Badino G (2010) Underground meteorology – "What's the weather underground?" Acta Carsologica 39(3): 427–448. https://doi.org/10.3986/ac.v39i3.74

Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: how, where and how many? Methods in Ecology and Evolution 3(2): 327–338. https://doi.org/10.1111/j.2041-210X.2011.00172.x

Barve N, Barve V, Jiménez-Valverde A, Lira-Noriega A, Maher SP, Peterson AT, Soberón J, Villalobos F (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecological Modelling 222(11): 1810–1819. https://doi.org/10.1016/j.ecolmodel.2011.02.011

Beaumont LJ, Hughes L, Pitman AJ (2008) Why is the choice of future climate scenarios for species distribution modelling important? Ecology Letters 11(11): 1135–1146. https://doi.org/10.1111/j.1461-0248.2008.01231.x

Bellard C, Bertelsmeier C, Leadley P, Thuiller W, Courchamp F (2012) Impacts of climate change on the future of biodiversity. Ecology Letters 15(4): 365–377. https://doi.org/10.1111/j.1461-0248.2011.01736.x

Breiner FT, Guisan A, Bergamini A, Nobis MP (2015) Overcoming limitations of modelling rare species by using ensembles of small models. Methods in Ecology and Evolution 6(10): 1210–1218. https://doi.org/10.1111/2041-210X.12403

Breiner FT, Nobis MP, Bergamini A, Guisan A (2018) Optimizing ensembles of small models for predicting the distribution of species with few occurrences. Methods in Ecology and Evolution 9(4): 802–808. https://doi.org/10.1111/2041-210X.12957

Cardoso P, Erwin TL, Borges PAV, New TR (2011) The seven impediments in invertebrate conservation and how to overcome them. Biological Conservation 144(11): 2647–2655. https://doi.org/10.1016/j.biocon.2011.07.024

Chen Z, Auler AS, Bakalowicz M, Drew D, Griger F, Hartmann J, Jiang G, Moosdorf N, Richts A, Stevanovic Z, Veni G, Goldscheider N (2017) The world karst aquifer mapping project: concept, mapping procedure and map of Europe. Hydrogeology Journal 25(3): 771–785. https://doi.org/10.1007/s10040-016-1519-3

China Nature Reserve Specimen Resource Sharing Platform (2024) List and vector boundaries of nature reserves in China Data set. Zenodo. https://doi.org/10.5281/zenodo.14875797 [accessed on 12 February 2026]

Christman MC, Culver DC, Madden MK, White D (2005) Patterns of endemism of the eastern North American cave fauna. Journal of Biogeography 32(8): 1441–1452. https://doi.org/10.1111/j.1365-2699.2005.01263.x

Culver DC, Pipan T (2019) The Biology of caves and other subterranean habitats (2nd edn.). Oxford University Press, Oxford, [xx +] 301 pp. https://doi.org/10.1093/oso/9780198820765.001.0001

Deharveng L, Bedos A (2012) Diversity patterns in the tropics. In: White WB, Culver DC (Eds) Encyclopedia of Caves (2nd edn.). Academic Press, New York/London, 238–250. https://doi.org/10.1016/B978-0-12-383832-2.00032-3

Di Cola V, Broennimann O, Petitpierre B, Breiner FT, D'Amen M, Randin C, Engler R, Pottier J, Pio D, Dubuis A, Pellissier L, Mateo RG, Hordijk W, Salamin N, Guisan A (2017) ecospat: an R package to support spatial analyses and modeling of species niches and distributions. Ecography 40(6): 774–787. https://doi.org/10.1111/ecog.02671

Didan K (2021) MODIS/Terra Vegetation Indices 16-Day L3 Global 1km SIN Grid V061. NASA Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MOD13A2.061 [accessed on 5 February 2026]

Dobrowski SZ (2011) A climatic basis for microrefugia: the influence of terrain on climate. Global Change Biology 17(2): 1022–1035. https://doi.org/10.1111/j.1365-2486.2010.02263.x

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber B, Lafourcade B, Leitão PJ, Münkemüller T, McClean C, Osborne PE, Reineking B, Schröder B, Skidmore AK, Zurell D, Lautenbach S (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36(1): 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Elith J, Graham C, Valavi R, Abegg M, Bruce C, Ford A, Guisan A, Hijmans RJ, Huettmann F, Lohmann L, Loiselle B, Moritz C, Overton J, Peterson AT, Phillips S, Richardson K, Williams S, Wiser SK, Wohlgemuth T, Zimmermann NE (2020) Presence-only and presence-absence data for comparing species distribution modeling methods. Biodiversity Informatics 15(2): 69–80. https://doi.org/10.17161/bi.v15i2.13384

Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40: 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. Methods in Ecology and Evolution 1(4): 330–342. https://doi.org/10.1111/j.2041-210X.2010.00036.x

Erickson KD, Smith AB (2023) Modeling the rarest of the rare: a comparison between multi-species distribution models, ensembles of small models, and single-species models at extremely low sample sizes. Ecography 2023(6): e06500. https://doi.org/10.1111/ecog.06500

Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, Taylor KE (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6): experimental design and organization. Geoscientific Model Development 9(5): 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Ficetola GF, Canedoli C, Stoch F (2019) The Racovitzan impediment and the hidden biodiversity of unexplored environments. Conservation Biology 33(1): 214–216. https://doi.org/10.1111/cobi.13179

Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37(12): 4302–4315. https://doi.org/10.1002/joc.5086

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24(1): 38–49. https://doi.org/10.1017/S0376892997000088

Fois M, Cuena-Lombraña A, Fenu G, Bacchetta G (2018) Using species distribution models at local scale to guide the search of poorly known species: review, methodological issues and future directions. Ecological Modelling 385: 124–132. https://doi.org/10.1016/j.ecolmodel.2018.07.018

Fourcade Y, Besnard AG, Secondi J (2018) Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. Global Ecology and Biogeography 27(2): 245–256. https://doi.org/10.1111/geb.12684

Franklin J (2010a) Mapping species distributions: spatial inference and prediction. Cambridge University Press, Cambridge, [xviii +] 320 pp. https://doi.org/10.1017/CBO9780511810602

Franklin J (2010b) Moving beyond static species distribution models in support of conservation biogeography. Diversity and Distributions 16(3): 321–330. https://doi.org/10.1111/j.1472-4642.2010.00641.x

Franklin J (2013) Species distribution models in conservation biogeography: developments and challenges. Diversity and Distributions 19(10): 1217–1223. https://doi.org/10.1111/ddi.12125

Graham CH, Elith J, Hijmans RJ, Guisan A, Peterson AT, Loiselle BA, The Nceas Predicting Species Distributions Working Group (2008) The influence of spatial errors in species

occurrence data used in distribution models. Journal of Applied Ecology 45(1): 239–247. https://doi.org/10.1111/j.1365-2664.2007.01408.x

Guéguen M, Blancheteau H, Lemaire-Patin R, Thuiller W (2026) biomod2: Ensemble platform for species distribution modeling. R package version 4.3-4-5. https://doi.org/10.32614/CRAN.package.biomod2

Guillera-Arroita G, Lahoz-Monfort JJ, Elith J, Gordon A, Kujala H, Lentini PE, McCarthy MA, Tingley R, Wintle BA (2015) Is my species distribution model fit for purpose? Matching data and models to applications. Global Ecology and Biogeography 24(3): 276–292. https://doi.org/10.1111/geb.12268

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecology Letters 8(9): 993–1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x

Guisan A, Tingley R, Baumgartner JB, Naujokaitis-Lewis I, Sutcliffe PR, Tulloch AIT, Regan TJ, Brotons L, McDonald-Madden E, Mantyka-Pringle C, Martin TG, Rhodes JR, Maggini R, Setterfield SA, Elith J, Schwartz MW, Wintle BA, Broennimann O, Austin M, Ferrier S, Kearney MR, Possingham HP, Buckley YM (2013) Predicting species distributions for conservation decisions. Ecology Letters 16(12): 1424–1435. https://doi.org/10.1111/ele.12189

Hao T, Elith J, Guillera-Arroita G, Lahoz-Monfort JJ (2019) A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. Diversity and Distributions 25(5): 839–852. https://doi.org/10.1111/ddi.12892

Hernandez PA, Graham CH, Master LL, Albert DL (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29(5): 773–785. https://doi.org/10.1111/j.0906-7590.2006.04700.x

Hijmans RJ, Brown A, Bivand MBR, Chirico M, Cordano E, Dyba K, Pebesma E, Rowlingson B, Sumner MD (2026) terra: spatial data analysis. R package version 1.8-93. https://doi.org/10.32614/CRAN.package.terra

Hirzel AH, Le Lay G, Helfer V, Randin C, Guisan A (2006) Evaluating the ability of habitat suitability models to predict species presences. Ecological Modelling 199(2): 142–152. https://doi.org/10.1016/j.ecolmodel.2006.05.017

Hortal J, de Bello F, Diniz-Filho JAF, Lewinsohn TM, Lobo JM, Ladle RJ (2015) Seven shortfalls that beset large-scale knowledge of biodiversity. Annual Review of Ecology, Evolution, and Systematics 46: 523–549. https://doi.org/10.1146/annurev-ecolsys-112414-054400

Jarnevich CS, Stohlgren TJ, Kumar S, Morisette JT, Holcombe TR (2015) Caveats for correlative species distribution modeling. Ecological Informatics 29(1): 6–15. https://doi.org/10.1016/j.ecoinf.2015.06.007

Jiang Z, Lian Y, Qin X (2014) Rocky desertification in Southwest China: impacts, causes, and restoration. Earth-Science Reviews 132: 1–12. https://doi.org/10.1016/j.earscirev.2014.01.005

Jiménez-Valverde A, Lobo JM, Hortal J (2008) Not as good as they seem: the importance of concepts in species distribution modelling. Diversity and Distributions 14(6): 885–890. https://doi.org/10.1111/j.1472-4642.2008.00496.x

Juberthie C, Decu V (1994) Encyclopaedia Biospeologica (Vol. I). Société de Biospéologie, Moulis, [iii–xii +] 1–834 pp.

Kass JM, Muscarella R, Galante PJ, Bohl CL, Pinilla-Buitrago GE, Boria RA, Soley-Guardia M, Anderson RP (2021) ENMeval 2.0: Redesigned for customizable and reproducible modeling of species' niches and distributions. Methods in Ecology and Evolution 12(9): 1602–1608. https://doi.org/10.1111/2041-210X.13628

Keppel G, Van Niel KP, Wardell-Johnson GW, Yates CJ, Byrne M, Mucina L, Schut AGT, Hopper SD, Franklin SE (2012) Refugia: identifying and understanding safe havens for biodiversity under climate change. Global Ecology and Biogeography 21(4): 393–404. https://doi.org/10.1111/j.1466-8238.2011.00686.x

Lee-Yaw JA, McCune JL, Pironon S, Sheth SN (2022) Species distribution models rarely predict the biology of real populations. Ecography 2022(6): e05877. https://doi.org/10.1111/ecog.05877

Leroy B (2023) Choosing presence-only species distribution models. Journal of Biogeography 50(1): 247–250. https://doi.org/10.1111/jbi.14505

Liu C, White M, Newell G (2013) Selecting thresholds for the prediction of species occurrence with presence-only data. Journal of Biogeography 40(4): 778–789. https://doi.org/10.1111/jbi.12058

Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography 17(2): 145–151. https://doi.org/10.1111/j.1466-8238.2007.00358.x

Lomba A, Pellissier L, Randin C, Vicente J, Moreira F, Honrado J, Guisan A (2010) Overcoming the rare species modelling paradox: A novel hierarchical framework applied to an Iberian endemic plant. Biological Conservation 143(11): 2647–2657. https://doi.org/10.1016/j.biocon.2010.07.007

Lumaret J-P, Lobo JM (1996) Geographic distribution of endemic dung beetles (Coleoptera, Scarabaeoidea) in the western Palaearctic Region. Biodiversity Letters 3(6): 192–199. https://doi.org/10.2307/2999676

Maitner BS, Richards RL, Carlson BS, Drake JM, Merow C (2026) Flexible methods for species distribution modeling with small samples. Ecography 2026(2): e08112. https://doi.org/10.1002/ecog.08112

Mammola S, Goodacre SL, Isaia M (2018) Climate change may drive cave spiders to extinction. Ecography 41(1): 233–243. https://doi.org/10.1111/ecog.02902

Mammola S, Leroy B (2018) Applying species distribution models to caves and other subterranean habitats. Ecography 41(7): 1194–1208. https://doi.org/10.1111/ecog.03464

Mammola S, Cardoso P, Culver DC, Deharveng L, Ferreira RL, Fišer C, Galassi DMP, Griebler C, Halse S, Humphreys WF, Isaia M, Malard F, Martinez A, Moldovan OT, Niemiller ML, Pavlek M, Reboleira ASPS, Souza-Silva M, Teeling EC, Wynne JJ, Zagmajster M (2019) Scientists' warning on the conservation of subterranean ecosystems. BioScience 69(8): 641–650. https://doi.org/10.1093/biosci/biz064

Mammola S, Amorim IR, Bichuette ME, Borges PAV, Cheeptham N, Cooper SJB, Culver DC, Deharveng L, Eme D, Ferreira RL, Fišer C, Fišer Ž, Fong DW, Griebler C, Jeffery WR, Jugovic J, Kowalko JE, Lilley TM, Malard F, Manenti R, Martínez A, Meierhofer MB, Niemiller ML, Northup DE, Pellegrini TG, Pipan T, Protas M, Reboleira ASPS, Venarsky MP, Wynne JJ, Zagmajster M, Cardoso P (2020) Fundamental research questions in subterranean biology. Biological Reviews 95(6): 1855–1872. https://doi.org/10.1111/brv.12642

Mammola S, Altermatt F, Alther R, Amorim IR, Băncilă RI, Borges PAV, Brad T, Brankovits D, Cardoso P, Cerasoli F, Chauveau CA, Delić T, Di Lorenzo T, Faille A, Fišer C, Flot J-F, Gabriel R, Galassi DMP, Garzoli L, Griebler C, Konecny-Dupré L, Martínez A, Mori N, Nanni V, Ogorelec Ž, Pallarés S, Salussolia A, Saccò M, Stoch F, Vaccarelli I, Zagmajster M, Zittra C, Meierhofer MB, Sánchez-Fernández D, Malard F (2024) Perspectives and pitfalls in preserving subterranean biodiversity through protected areas. npj Biodiversity 3: 2. https://doi.org/10.1038/s44185-023-00035-1

Mammola S, Meierhofer MB, Borges PA, Colado R, Culver DC, Deharveng L, Delić T, Di Lorenzo T, Dražina T, Ferreira RL, Fiasca B, Fišer C, Galassi DMP, Garzoli L, Gerovasileiou V, Griebler C, Halse S, Howarth FG, Isaia M, Johnson JS, Komerički A, Martínez A, Milano F, Moldovan OT, Nanni V, Nicolosi G, Niemiller ML, Pallarés S, Pavlek M, Piano E, Pipan T, Sanchez-Fernandez D, Santangeli A, Schmidt SI, Wynne JJ, Zagmajster M, Zakšek V, Cardoso P (2022) Towards evidence-based conservation of subterranean ecosystems. Biological Reviews 97(4): 1476–1510. https://doi.org/10.1111/brv.12851

McPherson JM, Jetz W (2007) Effects of species' ecology on the accuracy of distribution models. Ecography 30(1): 135–151. https://doi.org/10.1111/j.0906-7590.2007.04823.x

Mendes P, Velazco SJE, de Andrade AFA, Júnior PDM (2020) Dealing with overprediction in species distribution models: How adding distance constraints can improve model accuracy. Ecological Modelling 431: 109180. https://doi.org/10.1016/j.ecolmodel.2020.109180

Merow C, Smith MJ, Silander Jr JA (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography 36(10): 1058–1069. https://doi.org/10.1111/j.1600-0587.2013.07872.x

Merow C, Smith MJ, Edwards TC, Guisan A, McMahon SM, Normand S, Thuiller W, Wüest RO, Zimmermann NE, Elith J (2014) What do we gain from simplicity versus complexity in species distribution models? Ecography 37(12): 1267–1281. https://doi.org/10.1111/ecog.00845

Morales NS, Fernández IC, Baca-González V (2017) MaxEnt's parameter configuration and small samples: are we paying attention to recommendations? A systematic review. PeerJ 5: e3093. https://doi.org/10.7717/peerj.3093

Morelli TL, Barrows CW, Ramirez AR, Cartwright JM, Ackerly DD, Eaves TD, Ebersole JL, Krawchuk MA, Letcher BH, Mahalovich MF, Meigs GW, Michalak JL, Millar CI, Quiñones RM, Stralberg D, Thorne JH (2020) Climate-change refugia: biodiversity in the slow lane. Frontiers in Ecology and the Environment 18(5): 228–234. https://doi.org/10.1002/fee.2189

Moreno-Amat E, Mateo RG, Nieto-Lugilde D, Morueta-Holme N, Svenning J-C, García-Amorena I (2015) Impact of model complexity on cross-temporal transferability in Maxent species distribution models: An assessment using paleobotanical data. Ecological Modelling 312: 308–317. https://doi.org/10.1016/j.ecolmodel.2015.05.035

Moudrý V, Bazzichetto M, Remelgado R, Devillers R, Lenoir J, Mateo RG, Lembrechts JJ, Sillero N, Lecours V, Cord AF, Barták V, Balej P, Rocchini D, Torresani M, Arenas-Castro S, Man M, Prajzlerová D, Gdulová K, Prošek J, Marchetto E, Zarzo-Arias A, Gábor L, Leroy F, Martini M, Malavasi M, Cazzolla Gatti R, Wild J, Šímová P (2024) Optimising occurrence data in species distribution models: sample size, positional uncertainty, and sampling bias matter. Ecography 2024(12): e07294. https://doi.org/10.1111/ecog.07294

Nelder JA, Wedderburn RWM (1972) Generalized linear models. Journal of the Royal Statistical Society, Series A 135(3): 370–384. https://doi.org/10.2307/2344614

O'Neill BC, Kriegler E, Ebi KL, Kemp-Benedict E, Riahi K, Rothman DS, van Ruijven BJ, van Vuuren DP, Birkmann J, Kok K, Levy M, Solecki W (2017) The roads ahead: narratives for shared socioeconomic pathways describing world futures in the 21st century. Global Environmental Change 42: 169–180. https://doi.org/10.1016/j.gloenvcha.2015.01.004

Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeography 34(1): 102–117. https://doi.org/10.1111/j.1365-2699.2006.01594.x

Pebesma E (2018) Simple features for R: standardized support for spatial vector data. The R Journal 10: 439–446. https://doi.org/10.32614/RJ-2018-009

Pebesma E, Bivand R (2023) Spatial data science: with applications in R. Chapman and Hall/CRC, New York, 314 pp. https://doi.org/10.1201/9780429459016

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecological Modelling 190(3–4): 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A, Leathwick J, Ferrier S (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecological Applications 19(1): 181–197. https://doi.org/10.1890/07-2153.1

Phillips SJ, Anderson RP, Dudík M, Schapire RE, Blair ME (2017) Opening the black box: an open-source release of Maxent. Ecography 40(7): 887–893. https://doi.org/10.1111/ecog.03049

Qiao H, Soberón J, Peterson AT (2015) No silver bullets in correlative ecological niche modelling: insights from testing among many potential algorithms for niche estimation. Methods in Ecology and Evolution 6(10): 1126–1136. https://doi.org/10.1111/2041-210X.12397

Qiao H, Peterson AT, Ji L, Hu J (2017) Using data from related species to overcome spatial sampling bias and associated limitations in ecological niche modelling. Methods in Ecology and Evolution 8(12): 1804–1812. https://doi.org/10.1111/2041-210X.12832

R Core Team (2025) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org [accessed on 4 February 2026]

Raes N, ter Steege H (2007) A null-model for significance testing of presence-only species distribution models. Ecography 30(5): 727–736. https://doi.org/10.1111/j.2007.0906-7590.05041.x

Riahi K, van Vuuren DP, Kriegler E, Edmonds J, O'Neill BC, Fujimori S, Bauer N, Calvin K, Dellink R, Fricko O, Lutz W, Popp A, Cuaresma JC, KC S, Leimbach M, Jiang L, Kram T, Rao S, Emmerling J, Ebi K, Hasegawa T, Havlik P, Humpenöder F, Da Silva LA, Smith S, Stehfest E, Bosetti V, Eom J, Gernaat D, Masui T, Rogelj J, Strefler J, Drouet L, Krey V, Luderer G, Harmsen M, Takahashi K, Baumstark L, Doelman JC, Kainuma M, Klimont Z, Marangoni G, Lotze-Campen H, Obersteiner M, Tabeau A, Tavoni M (2017) The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emis-

sions implications: an overview. Global Environmental Change 42: 153–168. https://doi.org/10.1016/j.gloenvcha.2016.05.009

Sánchez-Fernández D, Galassi DMP, Wynne JJ, Cardoso P, Mammola S (2021) Don't forget subterranean ecosystems in climate change agendas. Nature Climate Change 11: 458–459. https://doi.org/10.1038/s41558-021-01057-y

Sillero N, Arenas-Castro S, Enriquez-Urzelai U, Vale CG, Sousa-Guedes D, Martínez-Freiría F, Real R, Barbosa A (2021) Want to model a species niche? A step-by-step guideline on correlative ecological niche modelling. Ecological Modelling 456: 109671. https://doi.org/10.1016/j.ecolmodel.2021.109671

Simon KS, Benfield EF, Macko SA (2003) Food web structure and the role of epilithic biofilms in cave streams. Ecology 84(9): 2395–2406. https://doi.org/10.1890/02-334

Sinclair SJ, White MD, Newell GR (2010) How useful are species distribution models for managing biodiversity under future climates? Ecology and Society 15(1): 8. https://doi.org/10.5751/es-03089-150108

Smith AB, Godsoe W, Rodríguez-Sánchez F, Wang H-H, Warren D (2019) Niche estimation above and below the species level. Trends in Ecology & Evolution 34(3): 260–273. https://doi.org/10.1016/j.tree.2018.10.012

Soley-Guardia M, Alvarado-Serrano DF, Anderson RP (2024) Top ten hazards to avoid when modeling species distributions: a didactic guide of assumptions, problems, and recommendations. Ecography 2024(4): e06852. https://doi.org/10.1111/ecog.06852

Sousa-Silva R, Alves P, Honrado J, Lomba A (2014) Improving the assessment and reporting on rare and endangered species through species distribution models. Global Ecology and Conservation 2: 226–237. https://doi.org/10.1016/j.gecco.2014.09.011

Stanton JC, Pearson RG, Horning N, Ersts P, Akçakaya HR (2012) Combining static and dynamic variables in species distribution models under climate change. Methods in Ecology and Evolution 3(2): 349–357. https://doi.org/10.1111/j.2041-210X.2011.00157.x

Swets JA (1988) Measuring the accuracy of diagnostic systems. Science 240(4857): 1285–1293. https://doi.org/10.1126/science.3287615

Tatebe H, Ogura T, Nitta T, Komuro Y, Ogochi K, Takemura T, Watanabe S, Kimoto M, Kawai H, Koshiro T, Oshima N, Yoshida K, Urakawa S, Tsujino H, Deushi M, Mishima M, Nosaka M, Saito R, Sasaki H, Koshiro T, Okajima H, Saito H, Suzuki T, Chikamoto Y, Ogura F, Aihara K (2019) Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. Geoscientific Model Development 12(7): 2727–2765. https://doi.org/10.5194/gmd-12-2727-2019

Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, Collingham YC, Erasmus BFN, de Siqueira MF, Grainger A, Hannah L, Hughes L, Huntley B, van Jaarsveld AS, Midgley GF, Miles L, Ortega-Huerta MA, Townsend Peterson A, Phillips OL, Williams SE (2004) Extinction risk from climate change. Nature 427: 145–148. https://doi.org/10.1038/nature02121

Thuiller W (2004) Patterns and uncertainties of species' range shifts under climate change. Global Change Biology 10(12): 2020–2027. https://doi.org/10.1111/j.1365-2486.2004.00859.x

Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD – a platform for ensemble forecasting of species distributions. Ecography 32(3): 369–373. https://doi.org/10.1111/j.1600-0587.2008.05742.x

Tian M-Y, Liu W-X, Huang S-B, Luo X-Z (2023) Cave biology. Geology Press, Beijing, [viii +] 248 pp. [in Chinese]

Tian M-Y, Wei G-F, Huang S-B, Liu W-X, Chen JJ, Li X, Wang C, Zheng Z, Shi K, Li A (2017) Life in dark world: cavernicolous creatures in Huanjiang World Natural Heritage site of South China Karst. Guangxi Science & Technology Publishing House, Nanning, [viii +] 134 pp. [in Chinese and English]

Tredennick AT, Hooker G, Ellner SP, Adler PB (2021) A practical guide to selecting models for exploration, inference, and prediction in ecology. Ecology 102(6): e03336. https://doi.org/10.1002/ecy.3336

Vaccarelli I, Colado R, Pallarés S, Galassi DMP, Sánchez-Fernández D, Di Cicco M, Meierhofer MB, Piano E, Di Lorenzo T, Mammola S (2023) A global meta-analysis reveals multilevel and context-dependent effects of climate change on subterranean ecosystems. One Earth 6(11): 1510–1522. https://doi.org/10.1016/j.oneear.2023.09.001

van Proosdij ASJ, Sosef MSS, Wieringa JJ, Raes N (2016) Minimum required number of specimen records to develop accurate species distribution models. Ecography 39(6): 542–552. https://doi.org/10.1111/ecog.01509

Velazco SJE, Rose MB, Marco PD, Regan HM, Franklin J (2024) How far can I extrapolate my species distribution model? Exploring shape, a novel method. Ecography 2024(3): e06992. https://doi.org/10.1111/ecog.06992

Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecological Applications 21(2): 335–342. https://doi.org/10.1890/10-1171.1

Wisz MS, Guisan A (2009) Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. BMC Ecology 9(1): 8. https://doi.org/10.1186/1472-6785-9-8

Wu T, Lu Y, Fang Y, Xin X, Li L, Li W, Jie W, Zhang J, Liu Y, Zhang L, Zhang F, Zhang Y, Wu F, Li J, Chu M, Wang Z, Shi X, Liu X, Wei M, Huang A, Zhang Y, Liu X (2019) The Beijing Climate Center Climate System Model (BCC-CSM): the main progress from CMIP5 to CMIP6. Geoscientific Model Development 12(4): 1573–1600. https://doi.org/10.5194/gmd-12-1573-2019

Yin Z-W (2020) New species of karst-dwelling Pselaphinae from southwestern China (Coleoptera: Staphylinidae). Acta Entomologica Musei Nationalis Pragae 60(1): 163–168. https://doi.org/10.37520/aemnp.2020.009

Yin Z-W (2025) Illuminating the darkness: an exceptionally diverse fauna of subterranean Pselaphinae (Coleoptera: Staphylinidae) in Asian karsts. Insect Systematics and Diversity 9(5): ixaf046. https://doi.org/10.1093/isd/ixaf046

Yin Z-W, He L (2020) New cavernicolous Pselaphinae from Sichuan, China (Coleoptera: Staphylinidae). The Coleopterists Bulletin 74(4): 827–836. https://doi.org/10.1649/0010-065x-74.4.827 065X-74.4.827

Yin Z-W, Li L-Z (2015) Zopherobatrus gen. n. (Coleoptera: Staphylinidae: Pselaphinae), a new troglobitic batrisine from southwestern China. Zootaxa 3985(2): 291–295. https://doi.org/10.11646/zootaxa.3985.2.8

Yukimoto S, Kawai H, Koshiro T, Oshima N, Yoshida K, Urakawa S, Tsujino H, Deushi M, Tanaka T, Hosaka M, Yabu S, Yoshimura H, Shindo E, Mizuta R, Obata A, Adachi Y, Ishii M (2019) The meteorological research institute earth system model version 2.0, MRI-ESM2.0: description and basic evaluation of the physical component. Journal of the Meteorological Society of Japan 97(5): 931–965. https://doi.org/10.2151/jmsj.2019-051

Zurell D, Franklin J, König C, Bouchet PJ, Dormann CF, Elith J, Fandos G, Feng X, Guillera-Arroita G, Guisan A, Lahoz-Monfort JJ, Leitão PJ, Park DS, Peterson AT, Rapacciuolo G, Schmatz DR, Schröder B, Serra-Diaz JM, Thuiller W, Yates KL, Zimmermann NE, Merow C (2020) A standard protocol for reporting species distribution models. Ecography 43(9): 1261–1277. https://doi.org/10.1111/ecog.04960

## Supplementary material 1

**ODMAP protocol describing the species distribution modelling used in this study, prepared following Zurell et al. (2020)**

Author: Zi-Wei Yin

Data type: docx

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/subtbiol.56.190752.suppl1

## Supplementary material 2

**Visualization of MaxEnt hyperparameter tuning results across 48 model configurations (*n* = 11)**

Author: Zi-Wei Yin

Data type: tif

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/subtbiol.56.190752.suppl2

## Supplementary material 3

**Model diagnostics for the Ensemble of Small Models (ESM)**
Author: Zi-Wei Yin
Data type: tif
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
Link: https://doi.org/10.3897/subtbiol.56.190752.suppl3

## Supplementary material 4

**Results of exhaustive MaxEnt hyperparameter tuning across 48 model configurations, evaluated using jackknife cross-validation**
Author: Zi-Wei Yin
Data type: docx
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
Link: https://doi.org/10.3897/subtbiol.56.190752.suppl4

## Supplementary material 5

**Summary of model evaluation for the Ensemble of Small Models (ESM) predicting the distribution of *Zopherobatrus***
Author: Zi-Wei Yin
Data type: docx
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
Link: https://doi.org/10.3897/subtbiol.56.190752.suppl5

## Supplementary material 6

**Sensitivity analysis comparing habitat retention projections between the full 4-variable ESM and a 3-variable ESM excluding NDVI**

Author: Zi-Wei Yin

Data type: docx

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/subtbiol.56.190752.suppl6